AN INTERACTIVE TEST OF MARINER COMPETENCE

EXECUTIVE SUMMARY

INTRODUCTION

The United States Coast Guard (USCG), in fulfillment of its responsibility for marine safety, provides for the examination and licensing of merchant mariners. At the present time, examinations are primarily pencil and paper tests of the candidate's <u>knowledge</u> of a topic. In a time of increased demand for testing of mariner competence, of decreased opportunity for formal on-board training, and of increased acceptance of simulators for training, the USCG sponsored an exploratory study of computer-based interactive testing of mariner competence. The expectation was that an interactive test would allow the examination of a candidate's <u>ability to apply knowledge</u> in a real-time decision-making context, an examination that is potentially better able to predict ability to perform successfully in the operational setting than is the current pencil and paper test. The purpose of our study was to explore the feasibility of interactive testing and its potential benefits, in order to determine whether the approach warranted further, more complete consideration.

OBJECTIVES AND SCOPE OF THE STUDY

The objectives of this study were, first, to explore the feasibility of <u>developing an interactive test</u>, using a desk-top simulator. The use of this type of technology could provide an interactive test that is potentially more accessible and affordable than is a full-mission simulator. Second, we explored the feasibility of <u>automatically scoring the interactive test</u>. Computer-based automatic scoring could free such testing from the requirement for expert examiners and provide objective, repeatable results. Finally, we explored the potential <u>results and benefits of an interactive test</u>. Test performance from a sample of future mariners provided a preview of the benefits to be expected from such assessment.

To limit the study to a workable scope, we explored only the assessment of a candidate's knowledge of, and ability to apply, the Rules of the Road (ROR). To do this, we used an existing desk-top simulator to provide a platform for our tester, an existing USCG Examination Module to define the initial test content, and a sample of United States Merchant Marine Academy (USMMA) cadets as a test population of future mariners. While the scope of the Interactive Rules of the Road Tester (IRORT) Project was limited, the issues considered are far broader and our exploration contributes to a general understanding of assessment of mariner competence by demonstration.

FEASIBILITY OF AN INTERACTIVE TEST

We selected the commercially-available desk-top simulator that best met our requirements for testing knowledge and application of ROR. This simulator presents the user with an "out the window" view which can be rotated 360 degrees and inspected with binoculars, bearing compass, and radar. All traffic ships have appropriate day shapes, lights, and whistle signals for their types. The mariner can control the course and speed of own ship and can sound whistle signals. For our purposes in designing the test, it was important that instructions and multiple-choice questions could be inserted into the scenarios and that a computer record was kept of all mariner and ship actions.

The test content was adapted from a Third Mate's Rules of the Road Examination Module. Our Subject Matter Experts (SME) examined the module and classified items as factual/objective, recognition, or operational. Each type of item was treated differently in the interactive test. Items deemed "factual/objective" by SMEs were presented on the computer in the same multiple-choice format as in the paper and pencil Module. "Recognition" items were treated by allowing the mariner to make observations of traffic ships in a dynamic context before answering embedded multiple-choice questions about them and the possible threat that they presented. "Operational" items, that required an understanding of the responsibilities to act imposed by ROR, were a tiny minority of the items in the paper and pencil Module, but were the focus of the interactive test. Three interactive scenarios were developed, to examine the mariners' ability to apply ROR under daytime, nighttime, and fog conditions.

FEASIBILITY OF AN AUTOMATIC SCORING SYSTEM

We designed an automatic scoring approach that replaced our SMEs' evaluations of mariners' performance with a procedure that could be applied by the computer. To capture the experts' judgments, we used an iterative process of review of each step of the scoring problem, independent input from each one on what was required, and group discussion until a consensus was reached. The agreed-upon basic testing objective was that the mariner be required to demonstrate an understanding of the requirements imposed by navigational law. These were to:

- maintain a good lookout and determine if risk of collision exists

- take appropriate action or maneuver to avoid collision

- determine if own ship's action or maneuver was adequate to avoid collision, and ensure that the action of maneuver does not put own ship in a close quarters situation with other vessels.

After agreeing on these objectives, the SMEs reviewed the three scenarios and selected specific operations that demonstrated understanding of each of these requirements. As an example, maintaining a lookout was demonstrated by appropriate visual search, inspection using binoculars, taking of bearings, and use of radar.

To set performance standards for each of these specific operations, our SMEs proposed the application of two scales: level of "proficiency" or level of "competency." Most of the operations were to be rated as to level of proficiency, which was defined with reference to both navigational law and professional standards as "expert," "qualified," or "unqualified." As an example, a mariner's visual search was expert, qualified, or unqualified as it compared to distributions of percent of time spent looking in each direction that were specified by the SMEs as expected for each scenario. Only a few of the operations were rated for level of competence, which was defined only with reference to navigational law. As an example, the mariner either met the legal requirement to sound a signal at maneuver, or he/she did not.

POTENTIAL RESULTS AND BENEFITS OF INTERACTIVE TESTS

Our final objective was the exploration of the potential results and benefits of an interactive test, compared to a multiple-choice paper and pencil test. Our technical approach to this exploration was to administer the test, designed by our SMEs, to a sample of 100 cadets at the U.S. Merchant Marine Academy. The administration both increased our understanding of the requirements for such a test and gave us a sample of performance data for analysis. For our analysis, we had three types of data for each cadet: performance on our interactive test, performance on a comparison multiple-choice paper and pencil test, and biographical data including grades in relevant courses. In addition, for the 50 First Classmen (seniors) in our sample, we had the grade for a course of full-mission exercises performed on a shiphandling simulator.

Our analysis of the data showed the following:

- The multiple-choice format (paper and pencil or computer-based) provided the highest correlation with our best measure of cadet <u>knowledge</u> -- scores in relevant classroom courses.

- The interactive operational scores provided the highest correlation with our best measure of cadet <u>application-based performance</u> -- scores in the full-mission exercises.

- The <u>combination</u> of multiple-choice and operational scores was needed to provide the highest correlation to the broad-based assessment of cadet capability provided by the full set of cadet biographical measures.

SUMMARY OF TECHNICAL CONCLUSIONS

An interactive Rules of the Road test is feasible, using a low fidelity desk-top simulator. Despite the obvious limitations to the realism of a desk-top system, our test required the cadets to demonstrate the ability to apply knowledge of ROR in a challenging, real-time situation.

Automatic scoring of an interactive test is feasible. Automatic scoring means that administration of the test requires a desk-top simulator and minimum attention from a proctor, but does not require an expert mariner to score each exercise history.

A meaningful assessment of the range of knowledge, skills, and abilities required to successfully fulfill the performance requirements imposed by ROR requires a combination of knowledge-based and application-based components.

SUMMARY OF RECOMMENDATIONS

We recommend further development of an interactive tester using a desk-top simulator for the assessment of ROR competence. Our recommended approach to this development is summarized in the report in Section 5. The approach is sufficiently general to apply to the assessment of other mariner competencies, whether using desk-top simulators, part-task trainers, real equipment, or other settings.